Aula
8

AMOSTRAGEM: DIMENSIONAMENTO DE AMOSTRAS. SELEÇÃO DOS ELEMENTOS DE UMA AMOSTRA. ESTIMATIVA DA CARACTERÍSTICA TOTAL DA POPULAÇÃO INVESTIGADA

META

Dimensionar o tamanho ideal de amostra para cada população. Estudo e Análise de populações com base nas características pesquisadas pela amostra.

OBJETIVOS

Ao final desta aula, o estudante deverá:
Saber dimensionar o tamanho ótimo de amostras.
Interpretar e analisar os resultados obtidos na amostra com conclusões sobre a população investigada.

PRÉ-REQUISITO:

Conhecimentos sobre estatística descritiva e probabilidades com ênfase para distribuição normal e todo conhecimento da aula anterior. Também são importantes: Papel, Calculadora ou Computador para realização dos cálculos.

INTRODUÇÃO

Olá! Tudo bem? Vamos dar seqüência ao estudo da estatística concluindo estudos sobre amostragem. Para isto é fundamental que você esteja em dia com os conceitos da estatística descritiva, principalmente os que envolvem cálculos sobre média aritmética e desvio padrão. Também são importantes os conhecimentos sobre probabilidade, com destaque para o uso e aplicação da distribuição normal.

Neste estudo final sobre amostragem é importante que você além dos conhecimentos acima mencionados, também esteja em dia com os temas tratados na aula anterior que é a parte inicial da montagem de uma pesquisa por amostragem.

Para que a pesquisa por amostragem tenha parâmetros confiáveis no estudo e análise da população investigada você deve ter muito cuidado em determinar o tamanho ótimo de amostra, que depende do nível de confiança da pesquisa, bem como do erro amostral, principalmente este último por está influenciado pela associação existente entre precisão da amostra, tempo de execução da pesquisa e recursos financeiros disponíveis.

Você tomará conhecimento como devem ser escolhidos os elementos de uma amostra e quais as principais metodologias a serem utilizadas. Como tudo que acontece com amostragem, esta também é uma etapa muito importante, visto que estes elementos devem ser selecionados de forma aleatória, isto é sem a interferência do pesquisador. Nesta escolha não pode haver preferências por esta ou aquela unidade da população, caso venha a ocorrer tal alternativa com certeza todo trabalho da pesquisa será prejudicado.

Aplicando adequadamente todos os critérios para montagem de uma pesquisa cientifica por amostragem você vai poder gerar a partir da amostra incadores populacionais consistentes e eficientes. Lembrando que o objetivo da amostra não é gerar uma estimativa exata do parâmetro populacional, mas que ele seja um bom estimador ao ponto de ser insignificante possível diferença entre eles.

DIMENSIONAMENTO DE AMOSTRAS

O tamanho da amostra além de depender do nível de confiança estabelecido para pesquisa é função direta do erro amostral, previamente fixado pelo pesquisador. O erro amostral mais utilizado é de 5%, vindo a seguir o de 1%. No máximo esse erro pode ser de 10%, para que se possa manter uma precisão razoável nos parâmetros estimados pela pesquisa. Quanto menor o erro amostral maior o tamanho da amostra, implicando em maior precisão nas estimativas populacionais e consequentemente maior custo para execução da pesquisa. A fixação do erro amostral dependa essencialmente de 3 fatores:

- Grau de precisão requerido pela pesquisa
- Grau de variabilidade da variável pesquisada
- Recursos financeiros disponíveis para execução da pesquisa.

Se a variável escolhida for quantitativa contínua, podemos determinar o tamanho da amostra, usando as fórmulas a seguirem:

População infinita: $n = \dfrac{z^2 \cdot \sigma^2}{E^2}$

População finita: $n = \dfrac{z^2 \cdot \sigma^2 \cdot N}{E^2 (N-1) + z^2 \cdot \sigma^2}$

Se a variável escolhida for quantitativa discreta, podemos determinar o tamanho da amostra, usando as fórmulas a seguir:

População infinita: $n = \dfrac{z^2 \cdot {'p} \cdot {'q}}{E_r^2}$

População finita: $n = \dfrac{z^2 \cdot {'p} \cdot {'q} \cdot N}{E_r^2 (N-1) + z^2 \cdot {'p} \cdot {'q}}$

z = Abscissa da curva normal, de acordo com o nível de confiança adotado.

σ^2 = variância da população. Caso seja desconhecido e não existam estudos semelhantes pode ser estimado a partir de uma amostra piloto em torno de 50 elementos, que geralmente é utilizada para testar o instrumento de pesquisa, bem como avaliar o comportamento da variável pesquisada.

E = erro amostral, expresso na unidade da variável pesquisada. Este erro representa a máxima diferença que o pesquisador admite existir entre μ e $\overline{x}$, isto é: $| \mu - \overline{x} | < E$. Bem como entre p e ‘p, isto é $| p - \text{‘}p | < E_r$.

Erro absoluto (E): É o produto do erro relativo (E_r) pela média aritmética da variável pesquisada. $E = E_r \cdot \overline{x}$

‘p = probabilidade de ocorrência da característica a ser observada. Quando este valor for desconhecido, se utiliza os dados da pesquisa piloto ou consideramos ‘p = 0,5 e dessa forma teremos o maior tamanho da amostra e consequentemente maior probabilidade de acerto. Observe que: p + q = 1.

SELEÇÃO DOS ELEMENTOS DA AMOSTRA

Os dois melhores processos para seleção dos elementos da amostra são: os números aleatórios e a seleção sistemática.

Números aleatórios: é constituída por conjuntos de números aleatórios, dispostos em colunas, podendo ser utilizados da esquerda para direita, de baixo para cima e vice-versa, desde que se mantenha uniformidade no método de seleção de cada amostra. Para melhor utilização das tabelas aleatórias, devemos trabalhar com o mesmo número de dígitos da população investigada. Para sua aplicação é preciso que todos os elementos da população estejam cadastrados para que possam ser identificados quando sorteados.
Seleção Sistemática – pode ser aplicada em qualquer população, mesmo que seus elementos não estejam cadastrados. Consiste no cálculo do intervalo de amostragem (I = N/n) de onde é selecionado, por qualquer processo, o primeiro elemento da amostra. Os demais são escolhidos, em intervalos constantes e iguais ao Intervalo de Amostragem, a partir do elemento anteriormente sorteado.

Exercícios resolvidos

1. Determinar o tamanho de uma amostra com precisão de 95%, para pesquisar a incidência de listeriose em uma população de 40.000 pessoas, considerando uma prevalência de anos anteriores igual a 30%. Usar erro amostral de 3%.

$$n = \frac{z^2 \text{ ‘}p \text{ ‘}q \, N}{E_r^2 (N - 1) + z^2 \text{ ‘}p \text{ ‘}q} = \frac{1{,}96^2 \, . \, 0{,}30 \, . \, 0{,}70 \, . \, 40.000}{0{,}03^2 (40.000 - 1) + 1{,}96^2 \, . \, 0{,}30 \, . \, 0{,}70} \Rightarrow$$

n = 877 pessoas.

2. Determinada variável apresentou, em relação ao seu peso médio, desvio padrão de 12 kg. Admitindo um nível de confiança de 95,5% e erro amostral de 1,2 kg, determine o tamanho da amostra para estudar essa população.

$$n = \frac{z^2\,\sigma^2}{E^2} = \frac{2^2 \,.\, 12^2}{1{,}2^2} \quad \Rightarrow \quad n = 400 \text{ peças}$$

De quanto deve ser o tamanho da amostra para uma população de 1.300 peças.

$$n = \frac{z^2\,\sigma^2\,N}{E^2\,(N-1) + z^2\,\sigma^2} = \frac{2^2 \,.\, 12^2 \,.\, 1.300}{12^2\,(1.300 - 1) + 2^2 \,.\, 12^2} \Rightarrow n = 306 \text{ peças.}$$

3. A proporção de eleitores favoráveis a determinado candidato é de 40%. Determinar o tamanho da amostra para investigar a tendência de voto desta população ao nível de confiança de 99% e erro amostral de 3%

$$n = \frac{z^2\;\text{‘}p\;\text{‘}q}{E_r^2} = \frac{2{,}57^2 \,.\, 0{,}40 \,.\, 0{,}60}{0{,}03^2} \quad \Rightarrow \quad n = 1.761$$

Determine o tamanho da amostra, admitindo uma população de 20.000 eleitores.

$$n = \frac{z^2\;\text{‘}p\;\text{‘}q\;N}{E_r^2\,(N-1) + z^2\;\text{‘}p\;\text{‘}q} = \frac{2{,}57^2 \,.\, 0{,}40 \,.\, 0{,}60 \,.20.000}{0{,}03^2\,(20.000 - 1) + 2{,}57^2 \,.\, 0{,}40 \,.\, 0{,}60} \Rightarrow$$

n = 1619 eleitores.

4. De um Aviário, que representa cerca de 1240 unidades, foi extraída, aleatoriamente, uma amostra de seis aves, cujos pesos acusaram os valores seguintes em kg: 2,3; 2,0; 2,4; 1,95; 1,85; 2,1. Determinar as estimativas da média e variância populacional

$$\bar{x} = \frac{\sum x_i}{n} = \frac{2{,}3 + 2{,}0 + 2{,}4 + 1{,}95 + 1{,}85 + 2{,}1}{6} = \frac{12{,}6}{6} = 2{,}1\text{kg}$$

$$E\,[\bar{x}] = \mu \quad \Rightarrow \quad \mu = 2{,}1\text{kg}$$

$$s^2 = \frac{2{,}3^2 + 2{,}0^2 + 2{,}4^2 + 1{,}95^2 + 1{,}85^2 + 2{,}1^2}{6} - 2{,}1^2 = 0{,}0375\text{kg}$$

$$\sigma_{\bar{x}}^2 = \frac{n}{n-1} \,.\, \frac{N-1}{N} \,.\, s^2 = \frac{6 \,.\, 1239 \,.\, 0{,}0375}{5 \quad 1240} = \frac{278{,}775}{6200} = 0{,}04996\text{ kg}^2$$

Observe-se que se tivéssemos usado a fórmula da população infinita, o valor de σ^2 pouco iria diferir do que foi encontrado, pois, N é muito grande.

ATIVIDADES

1. Um conjunto de 25 animais apresentou os seguintes pesos: 30, 22, 32, 22, 30, 28, 40, 32, 36, 32, 33, 28, 30, 32, 38, 40, 35, 40, 22, 40, 22, 45, 45, 38 e 48. Estimar uma amostra para esta população com precisão de 95% e erro amostral de 5%.

2. Uma amostra aleatória 400 estudantes de uma Universidade apresentou 96 deles devendo disciplinas em seus currículos. Adotando-se o nível de confiança de 95%, que tamanho deve ter a amostra para que o erro máximo de estimativas não exceda 2%?

3. Em um aviário foram selecionados ao acaso cerca de 10 frangos, os quais passaram a ser tratados, durante algumas semanas, com um tipo especial de ração, observando-se os seguintes aumentos de peso, em gramas: 250, 400, 200, 320, 250, 380, 190, 240, 390 e 280. Com precisão de 95% e erro de 5% que tamanho de amostra deveria tomar desta população para estudo e análise da referida característica?

4. Examinando 200 lâmpadas produzidas por uma empresa, foram encontradas 5 defeituosas. Adotando o nível de confiança de 95% :
a) Estimar p;
b) Qual o tamanho da amostra que possibilitaria um erro máximo de estimativa de 1% ?

5. Numa pesquisa de mercado bem conduzida, 57 dentre 150 entrevistados afirmaram que seriam compradores de certo produto a ser lançado. Sendo a população de compradores em potencial formada por 2000 elementos, qual o número mínimo de pessoas que comprarão o produto, com precisão de 95%.

6. Com a finalidade de se estudar a influência da determinada dieta no ganho de peso de prematuros, foi observada uma amostra piloto de 33 recém-nascidos (com peso entre 1500 e 2000 g). Estime o peso médio dessa população com precisão de 95% e dimensione qual deveria ser o tamanho de amostra, com erro amostral de 5%, para investigar essa população, considerando um total de 1.400 crianças nessa faixa de peso.

22	22	31	31	26	27	22	26	24	21	20
42	42	30	28	26	25	27	24	23	25	20
40	31	28	27	42	32	24	29	23	24	30

7. Uma amostra simples ao acaso de 30 domicílios foi selecionada de uma zona urbana que contém 15000 domicílios. O numero de pessoas de cada um dos domicílios que integram a amostra é o seguinte:

5	6	3	3	2	3	3	3	4	4	3	2	7	4	3
5	4	4	3	3	4	3	3	1	2	4	3	4	2	4

Estimar o numero total de pessoas que vivem nesta área. Utilize o método de estimativa pontual e por intervalo.

CONCLUSÃO

A partir desta aula você deve saber como determinar o tamanho ótimo de amostra para investigar uma população, isto é: aquela amostra dimensionada com base no nível de confiança e erro amostral, para que você tenha conclusões precisas acerca da natureza da população investigada.

A partir desta técnica de pesquisa pode investigar determinado caráter de uma população quando só conhece dados de uma amostra, por isto que a escolha do tamanho ótimo é tão importante. Esse é um problema que vai enfrentar quando, por exemplo, quer estudar a relação entre o fumo e o câncer de pulmão a partir de uma amostra, visto que as conclusões deste estudo é inferir os resultados da amostra para todos os indivíduos da população.

Tão importante como o tamanho da amostra é a seleção dos seus elementos. Alguns métodos de seleção dos elementos de uma amostra foram indicados e entre eles os mais utilizados são aqueles que trabalham com números aleatórios e a seleção sistemática. Você já aprendeu que a seleção aleatória pode ser aplicada quando for possível obter um cadastro da população investigada, enquanto a seleção sistemática é mais apropriada para aquelas populações que não possuem cadastro dos seus elementos, embora nada impeça que ela seja aplicada também em populações que tenham cadastro.

No final da aula temos uma lista de exercícios para serem resolvidos em grupos de no máximo cinco pessoas ou individual. Com certeza você vai ficar muito satisfeito com os resultados do seu desempenho.

RESUMO

Nesta aula falamos principalmente do tamanho da amostra, seleção dos seus elementos e estimativas de características populacionais.

Mostramos que existem vários métodos para extrair uma amostra, sendo que vamos concentrar nossa atenção na amostragem aleatória simples, método este onde cada elemento da população tem a mesma oportunidade de ser incluído na amostra.

Foi estudado também que o tamanho de uma amostra além de depender do nível de confiança estabelecido para pesquisa é função direta do erro amostral, previamente fixado pelo pesquisador, isto é: quanto menor o erro maior o tamanho da amostra, implicando em maior precisão nas estimativas populacionais. Esta é uma situação ótima e ideal para qualquer pesquisa por amostragem, mas em contrapartida a precisão da pesquisa é função direta do tempo e dos recursos financeiros destinados para sua execução. Conclusão: nem sempre você pode trabalhar com a precisão ideal em decorrência da existência de algum limite nos outros fatores que também influenciam na escolha do tamanho do erro. Erro este que no máximo deve ser de 10%, sendo 5% e 1% os mais indicados.

Também escrevemos sobre os principais métodos de seleção dos elementos de uma amostra: números aleatórios e seleção sistemática. A tabela de números aleatórios é constituída por conjuntos de números escolhidos ao acaso, e dispostos em colunas. Para melhor utilização das tabelas aleatórias, devemos trabalhar com o mesmo número de dígitos da população investigada e sua aplicação depende de que os elementos da população estejam cadastrados.

A Seleção Sistemática – pode ser aplicada em qualquer população. Consiste no cálculo do intervalo de amostragem de onde é selecionado, por qualquer processo, o primeiro elemento da amostra. Os demais são escolhidos, em intervalos constantes e iguais ao Intervalo de Amostragem.

AUTO-AVALIAÇÃO

Sou capaz de determinar o tamanho ótimo de amostras para estudar populações?

Sou capaz de selecionar os elementos de uma amostra?

Sou capaz de interpretar e analisar as características de populações com base nas informações obtidas por amostras?

PRÓXIMA AULA

Testes de Hipóteses envolvendo o estudo dos testes: Normal, "t" de Student e Qui-Quadrado.

REFERÊNCIAS

RODRIGUES, PEDRO CARVALHO. **Bioestatística**. Universidade Federal Fluminense.
FONSECA, JAIRO DA. **Curso de Estatística**. Editora Atlas.
OLIVEIRA, FRANCISCO ESTEVAM MARTINS DE. **Estatística e Probabilidade**. Editora Atlas.
TANAKA. **Elementos de Estatística**. Editora McGraw.Hill.
BARBETTA, PEDRO A. **Estatística aplicada às Ciências Sociais**. Editora da UFSC.
GÓES, LUIZ A. C. **– Estatística I e II –** Editora Saraiva.
DÍAZ, FRANCISCA; LOPES, FRANCISCO JAVIER. **Bioestatística**. Editora Thomson.